引文格式：王小攀, 刘小立, 马泽忠, 等. 基于“珠海一号”的水稻高光谱定量估产研究[J]. 地理空间信息, 2023, 21(8): 76-79. doi:10. 3969/j. issn. 1672-4623.2023.08.019

基于“珠海一号”的水稻高光谱定量估产研究

王小攀[1]，刘小立[2]，马泽忠[1]，何　宗[1]，罗　鼎[1]
（1. 重庆市地理信息和遥感应用中心，重庆 401147；2. 重庆市规划和自然资源局机关事务中心，重庆 401147）

摘　要：利用“珠海一号”高光谱卫星遥感数据实现水稻定量估产反演。首先抽稀地面实测光谱，与卫星数据光谱波段保持一致；其次通过皮尔逊相关分析，从光谱、光谱一阶导数以及NDVI、RVI、DVI、EVI2、VI、SAVI等6个植被指数中，计算并选取产量敏感性较高的18个建模因子；再次通过训练随机森林模型，获得最优训练结果下的最优参数模型；最后利用珠海一号卫星数据进行定量反演，得到研究区的水稻估产值。结果表明：基于4参数的多生育期随机森林估产模型精度较高，抽穗期光谱704 nm波段一阶导，灌浆期567 nm波段一阶导，灌浆期884 nm波段一阶导以及灌浆期DVI（λ_1=658 nm，λ_2=523 nm）为自变量的模型效果最佳；珠海一号水稻反演效果较好，反演得到铁桥村水稻平均产量575.2 kg/亩，与实际采集情况相差仅15.5 kg/亩。

关键词：水稻；珠海一号；高光谱；随机森林；遥感
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2023）08-0076-04

Rice Yield Estimation Based on OHS-1 Hyperspectral Remote Sensing Images

WANG Xiaopan[1], LIU Xiaoli[2], MA Zezhong[1], HE Zong[1], LUO Ding[1]
(1. Chongqing Geomatics and Remote Sensing Application Center, Chongqing 401147, China;
2. Chongqing Planning and Natural Resources Bureau Offices Administration Center, Chongqing 401147, China)

Abstract: We focused on rice yield estimation based on OHS-1 hyperspectral remote sensing images. Firstly, we diluted the measured spectra on the ground and maintain consistency with the spectral bands of satellite data.Then,we calculated the eighteen modeling factors by Pirsson correlation analysis, from the spectra, the first derivative of spectra and six vegetation indices (NDVI, RVI, DVI, EVI2, VI and SAVI), and trained the random forest model to obtain the optimal parameter. Finally, we carried out the quantitative inversion by OHS-1to obtain the rice yield. The results show that the rice yield estimation of multi-growth period random forest based on four parameters has high accuracy,with the first derivative at 704 nm at heading stage and 567 nm at filling stage, the first-order conduction at 884 nm and DVI at grain filling stage (λ_1=658 nm, λ_2=523 nm) is the best independent variables. The average yield of TieqiaoVillage rice is 575.2 kg/mu, with the only 15.5 kg/mu difference from the truth.

Keyword: rice, OHS-1, hyperspectral, random forest, remote sensing

水稻产量高低和品质好坏对粮食产量安全有重要影响，利用高光谱遥感可以定量反演作物长势相关信息，能够为农作物保供保产从定性式、粗放式管理模式转变到定量化、精细化、高质量管理提供指导和服务[1]。

王延颐[2]等最早将高光谱遥感技术应用在水稻生产管理上，构建了基于不同生育期的光谱植被指数对南方稻区的水稻进行了估产。随着遥感技术的发展，Sentinel-2、MODIS、HJ-1A等国内外高光谱遥感影像数据在水稻估产中得到广泛应用[3-6]。近年来无人机技术兴起，机载高光谱逐渐应用到水稻估产中[4]。胡景辉[5]等利用近地面无人机平台获取的成像高光谱数据开展基于传统经验统计模型的多生育期水稻估产研究，洪雪[6]利用无人机高光谱遥感技术获取的光谱反演的水稻冠层归一化植被指数（normalized difference vegetation index，NDVI）和叶片光化学指数（photochemical reflectance index，PRI）数据进行估产建模。“珠海一号”作为我国唯一商业高光谱卫星遥感数据，至2018年首次组网运行以来，在叶绿素反演[7]、水质反演[8]以及农作物精细分类[9]等方面发挥了重要作用，但鲜有水稻产量估算相关的研究。

1　材料与方法

1.1　研究区概况

本次实验区位于重庆市南川区大观镇铁桥村，地处106.995 4°E、29.242 1N°，海拔高度为727 m，面积约1 500亩，按照均匀分布原则，共设计32个采样

收稿日期：2023-03-21；修回日期：2023-05-08。
项目来源：重庆市科研机构绩效激励引导专项（cstc2021jxjl00024）。
第一作者简介：王小攀（1989—），硕士，高级工程师，研究方向为农业遥感、农业高光谱遥感定量研究等，E-mail：wangxpcq@163.com。

点，样点分布如图1所示。



图1　水稻光谱测量采样点分布图

1.2　数据获取

1.2.1　地物光谱数据

水稻叶片及冠层高光谱数据采集使用的是美国ASD Filed Spec背挂式野外光谱仪，仪器采集的光谱范围为350~2 500 nm，主要采集水稻孕穗期、抽穗期、灌浆期、成熟期等4个生育期的光谱数据。地物测定前后，进行参考板校正，准同步地测量参考板反射的太阳辐射光谱，通过比值方法，计算目标的光谱反射率。反射率计算方式如式（1）所示。

$$R_T^n = \frac{DN_T^n}{DN_R^n} \times R_R^n \tag{1}$$

式中，R_T^n 和 R_R^n 分别为目标和参考板在第 n 个波段的光谱反射率；DN_T^n 和 DN_R^n 为目标和参考板在第 n 个波段反射的太阳光强度。

1.2.2　珠海一号高光谱卫星数据

本研究采用“珠海一号”高光谱卫星遥感数据。“珠海一号”高光谱卫星在轨8颗，均采用推扫成像方式，单次成像范围150 km×400 km/min，空间分辨率10 m，光谱分辨率2.5 nm，波长范围400~1 000 nm。本次实验接收了OHS-3D CMOS2、OHS-2D CMOS2、OHS-2A CMOS2、OHS-3A CMOS2四颗卫星传感器的拍摄数据。高光谱卫星遥感影像数据采集时间具体如表1所示。“珠海一号”高光谱数据处理使用ENVI 5.0，首先利用ENVI5.0中Flash大气校正模块进行辐射校正，再利用经过精校正的SPOT6影像进行几何校正，获得高光谱反射率数据。

表1　遥感影像和同步情况

生育期	影像采集时间	卫星传感器	地面光谱采集时间	采样时差
孕穗期Booting	2022-07-14	OHS-3D CMOS2	2022-07-15	晚1d
抽穗期Heading	2022-07-25	OHS-2D CMOS2	2022-07-26	晚1d
灌浆期Filling	2022-08-10	OHS-2A CMOS2	2022-08-10	同步
成熟期Ripening	2022-08-22	OHS-3A CMOS2	2022-08-22	同步

1.2.3　水稻产量数据

记录每个样点周围1 m²范围的单位面积内种植水稻的平均穴数。水稻收割前，在每个样区各个采样点1 m²范围内，随机抽选3穴，记录每穴水稻穗数，再在随机抽选的水稻穴中，随机采集3个穗，及时对采集的样本进行晒干处理，并测量称重。最后按照式（2）所示计算水稻理论产量。

水稻理论产量（单位）= 单位面积穴数×每穴平均有效穗数×平均穗粒数×结实率×百粒重　（2）

对采集的水稻产量数据进行了初步分析和筛选，扣除产量数据缺失、产量数据存在粗差的样本点，利用筛选后的27个样本点产量数据进行数据分析及模型构建工作。27个样本点产量数据中，最大产量为669.6 kg/亩，最小产量为335.8 kg/亩，平均产量为559. 7 kg/亩。

2　模型构建及评价

2.1　随机森林模型方法

随机森林算法是Bagging集成学习理论与随机子空间方法的结合，是一种优异的监督学习机器学习算法，在遥感信息提取方面运用较多[10]。随机森林算法分类决策可用如下式子表示：

$$H(x) = \arg\max_{Y} \sum_{i=1}^{k} I(h_i(x) = Y) \tag{3}$$

式中，$H(x)$ 为组合分类模型；h_i 为单个据册数分类模型；Y 为输出变量；$I(.)$为示性函数。

2.2　建模因子筛选

2.2.1　高光谱特征参数

选择高光谱特征植被指数以及光谱原始反射率、反射率一阶导来计算光谱特征参数。具体计算方法如表2所示。

表2　光谱特征参数计算方法

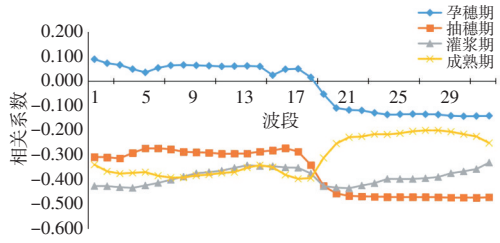
特征参数	计算公式
归一化植被指数(NDVI)	$(B_{NIR}-B_R)/(B_{NIR}+B_R)$
增强植被指数(EVI2)	$2.5(B_{NIR}-B_R)/(B_{NIR}+2.4_{BR}+1)$
比值植被指数(RVI)	B_{NIR}/B_R
差值植被指数(DVI)	$(B_{NIR}-B_R)$
植被指数(VI)	B_{NIR}/B_R-1
土壤调节植被指数(SAVI)	$(1+L)(B_{NIR}-B_R)/(B_{NIR}+2.4B_R+L)(L=0.5)$
光谱反射率 $\rho(\lambda_i)$	$\rho(\lambda_i)$
光谱反射率一阶导 $\rho'(\lambda_i)$	$\rho(\lambda_{i+1})-\rho(\lambda_{i-1}))/2\Delta\lambda$

2.2.2　特征参数敏感性分析

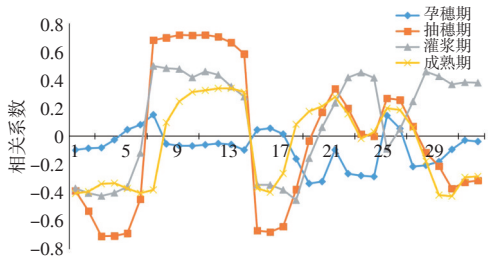
利用皮尔逊相关系数计算光谱原始反射率、反射

率一阶导数以及高光谱特征植被指数与产量数据的相关性。

1）光谱特征计算。产量与冠层光谱、冠层光谱一阶导数的相关系数计算结果如图2所示。通过比较相关系数，可以看出：一是从各生长期的冠层光谱与产量的相关性计算结果来看抽穗期相关性最高，孕穗期相关性较低。二是从各生长期的冠层光谱一阶导数与产量的相关性计算结果来看，抽穗期相关性最高，孕穗期相关性较低。



a 冠层光谱



b 冠层光谱一阶导数

图2 产量与光谱特征相关系数

2）植被指数特征计算。采用穷举法将32个波段两两组合分别计算不同波段组合下的NDVI、RVI、DVI、EVI2、VI、SAVI。将所有关键生育期所有波段组合计算出的植被指数与产量数据进行相关性分析，选取皮尔逊相关系数最高的波段组合作为相应生育期各植被指数的最佳敏感波段。可以看出：一是孕穗期、抽穗期植被指数与水稻产量相关性介于0.2~0.5之间，表现为较弱的相关性。二是灌浆期、成熟期植被指数与水稻相关性介于0.5~0.7之间，表现为较强的相关性。

2.3 建模因子

选取光谱、光谱一阶导数与产量相关性较好的10个光谱特征参数以及灌浆期DVI、RVI、NDVI、EVI2、VI，成熟期DVI、EVI2、SAVI等与水稻产量相关系数大于0.5，且显著性水平为0.01的8个植被指数，共计18个特征参数作为建模因子进行水稻估产建模。建模备选因子如表3所示。

2.4 模型训练

自变量Depth、mTry、nTree是随机森林模型重要参数，通过实验模拟发现随着Depth的变化，模型精度变化不大，随着nTree的变化，模型精度变化较大，在nTree=6~9时模型趋于稳定。因此模型训练设定nTree=6，Depth=100。mTry是输入的参数数量。我们首先选择了18个自变量作为模型构建的初始变量。并对18个因子进行自动筛选，以降低模型的复杂度。我们使用穷举法进行自变量筛选，设置模型输入自变量数量从1~18，比较模型训练的袋外正确率（OOB）大小，保留OOB值较大的模型，并获取最佳参数。比较不同自变量下模型最佳精度训练效果，其中训练效果最优的5种模型如表4所示。

表3 预选的建模因子

建模因子	释义	相关系数
2-Band24	抽穗期波段24（λ=826 nm）	−0.471 *
2-Band31	抽穗期波段31（λ=930 nm）	−0.474 *
2-Band3′	抽穗期波段3的一阶导数（λ=501 nm）	−0.711 **
2-Band9′	抽穗期波段9的一阶导数（λ=598 nm）	0.717 **
2-Band16′	抽穗期波段16一阶导数（λ=704 nm）	−0.681 **
3-Band4	灌浆期波段4（λ=523 nm）	−0.434 *
3-Band21	灌浆期波段21（λ=778 nm）	−0.436 *
3-Band7′	灌浆期波段7的一阶导数（λ=567 nm）	0.497 **
3-Band18′	灌浆期波段18一阶导数（λ=733nm）	−0.455 *
3-Band28′	灌浆期波段28一阶导数（λ=884 nm）	0.457 *
3-DVI	灌浆期DVI（λ_1=658 nm，λ_2=523 nm）	0.607 4**
3-RVI	灌浆期RVI（λ_1=838 nm，λ_2=778 nm）	0.536 3**
3-NDVI	灌浆期NDVI（λ_1=838 nm，λ_2=778 nm）	0.532 7**
3-EVI2	灌浆期EVI2（λ_1=658 nm，λ_2=523 nm）	0.599 **
3-VI	灌浆期VI（λ_1=838 nm，λ_2=778 nm）	0.536 3**
4-DVI	成熟期DVI（λ_1=899 nm，λ_2=913 nm）	0.504 **
4-EVI2	成熟期EVI2（λ_1=899 nm，λ_2=913 nm）	0.501 **
4-SAVI	成熟期SAVI（λ_1=899 nm，λ_2=913 nm）	0.496 2**

注：*表示在0.05水平下显著相关，**表示在0.01水平下显著相关。

表4 生育期耦合估产模型（RF-Multi）训练结果

模型编号	mTry	R^2	RMSE	OOB
RF-Multi1	4	0.951	21.671	0.637
RF-Multi2	5	0.946	22.777	0.608
RF-Multi3	6	0.942	23.576	0.585
RF-Multi4	7	0.945	22.815	0.573
RF-Multi5	8	0.937	24.451	0.586

从训练结果来看，自变量为′2-Band16’′，′3-Band7’′，′3-Band28’′，′3-DVI′，nTree=6，Depth=100时多生育期耦合模型表现最佳，其中R^2=0.951，RMSE=21.671 kg/亩。

2.5 精度验证

采用“留一法”对训练模型进行精度交叉验证，以任意1个样本为测试数据，其余26个样本为模型训练数据，最终得到27组预测值及模型精度测试结果。测试模型精度结果如表5所示。

表5　模型精度评价

评价指标	模型精度	评价指标	模型精度/(kg/亩)	评价指标	模型精度/%
R^2 最大值	0.961	RMSE最大值	22.08	平均绝对相对误差最大值	0.44
R^2 最小值	0.943	RMSE最小值	19.63	平均绝对相对误差最小值	0.01
R^2 均值	0.95	RMSE均值	21.66	平均绝对相对误差均值	0.12
R^2 方差	0.01	RMSE方差	0.42	平均绝对相对误差方差	0.01

将RF-Multi1模型预测值与亩产实测值绘制散点图，结果如图3所示。其中RF-Multi模型产量估算结果满足关系式：$y = 0.860\,6x + 76.258$，$R^2 = 0.961\,6$，表明模型预测效果较好。

2.6　产量估算

以南川区铁桥村的水田矢量底图，参考高分辨率遥感影像，通过人工目视解译进一步提取铁桥村水稻田范围，作为水稻产量反演的研究区域。基于水稻田矢量图层掩膜“珠海一号”影像数据，计算目标光谱参数，使用RF-Multi1模型进行产量估算。估算结果如图4所示，铁桥村水稻产量估计值范围为363.3~621.2 kg/亩，实际产量均值为550.7 kg/亩，全村水稻平均产量预计为575.2 kg/亩。

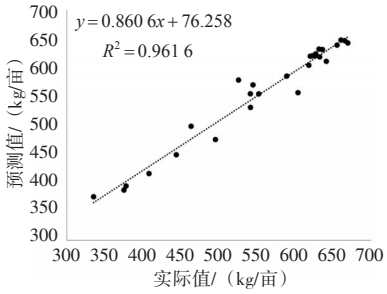


图3　RF-Multi1模型预测值与亩产实测值散点图

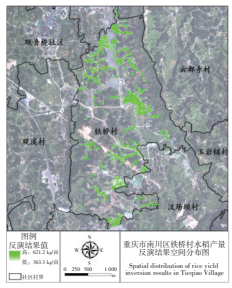


图4　铁桥村水稻产量反演结果空间分布图

3　结论和讨论

3.1　结　论

实验结果表明，一是参数多生育期随机森林估产模型精度较高，抽穗期光谱704 nm波段一阶导，灌浆期567 nm波段一阶导，灌浆期884 nm波段一阶导以及灌浆期DVI（λ_1=658 nm，λ_2=523 nm）为自变量的模型效果最佳。二是随着Depth的变化，模型精度变化不大，随着nTree的变化，模型精度变化较大，在nTree=6~9时模型趋于稳定，nTree=6，Depth=100效果最佳。三是在南川大观镇实验区进行定量反演，获得铁桥村水稻产量范围为363.3~621.2 kg/亩，全村水稻平均产量预计为575.2 kg/亩，产量预测结果与实际采集情况仅相差15.5 kg/亩，情况基本相符。

3.2　讨　论

本研究首次将空间分辨率10 m的“珠海一号”国产高光谱卫星遥感数据用于山地区域水稻产量估算工作，并取得较好的训练结果。基于多生育期多特征耦合的随机森林算法充分挖掘了不同生育期水稻光谱特征，尤其是抽穗期、灌浆期光谱，能够获得更佳的估算效果，为大范围水稻估产工作提供了可能。当然，也还有些困难需要进一步解决：一是研究中采集的ASD地面光谱数据与“珠海一号”高光谱遥感影像波段范围、采样间隔、反射率值等存在较大差异，利用地面光谱建模卫星影像反演的方式，可能会导致一定误差。二是利用遥感数据进行水稻估产，还缺少关键性的高精度水稻分布数据。10 m分辨率“珠海一号”数据在自动提取水稻种植范围精度方面，还无法满足农业部门管理工作需要，需要进一步协同高分辨率遥感影像提取空间分布图。

参考文献

[1] 肖天豪,周学军,李存文,等. 基于无人机高光谱的冬小麦植株氮含量估算[J]. 地理空间信息,2022,20(8):8-12

[2] 王延颐. 南方稻区遥感水稻长势监测与估产研究[J]. 遥感技术与应用,1991,6(3):1-6

[3] 吴燕平,洪勇,姜益民,等. 基于Sentinel-2和MODIS数据的宣恩县水稻遥感估产研究[J]. 测绘与空间地理信息,2022,45(12):33-37

[4] 王飞龙. 基于卫星和无人机多源数据的水稻遥感估产方法研究[D]. 杭州:浙江大学,2019

[5] 胡景辉. 基于无人机成像高光谱遥感数据的水稻估产方法研究[D]. 杭州:浙江大学,2020

[6] 洪雪. 基于水稻高光谱遥感数据的植被指数产量模型研究[D]. 沈阳:沈阳农业大学,2017

[7] 冯天时,庞治国,江威. 基于珠海一号高光谱卫星的巢湖叶绿素a浓度反演[J]. 光谱学与光谱分析,2022,42(8):2 642-2 648

[8] 殷子瑶,李俊生,范海生,等. 珠海一号高光谱卫星的于桥水库水质参数反演初步研究[J]. 光谱学与光谱分析,2021,41(2):494-498

[9] 张信耶,李继旭. 珠海一号高光谱卫星在高密市农作物精细分类中的应用[J]. 卫星应用,2019(8):29-33

[10] 赵宇虹. 随机森林遥感信息提取研究进展及应用研究[J]. 测绘与空间地理信息,2021,44(3):133-136